IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

United States Patent Application for

Machine Vision Method and Apparatus for Thresholding Images of Non-Uniform Materials

Inventors:

Jeffrey Michael Wolinsky,
2205 McGee Avenue, Berkeley, CA 94703,
a citizen of the United States and

Sherrill E. Lavagnino,
3362 Herrier Street, Oakland, CA 94602
a citizen of the United States

Field of the Invention

The present invention relates generally to machine vision and, more particularly, to thresholding images. The invention has application, for example, in imaging and analyzing

5   non-woven and other materials with high-frequency gray-scale variation, e.g., for irregularities and defects.

Background of the Invention

10   In machine vision, image enhancement techniques are used to process image data to facilitate operator and automated analysis. Commonly known image enhancement techniques can be divided into two broad classes: point transforms and neighborhood operations. Point transform algorithms are ones in which each output pixel is generated as a function of a corresponding input pixel. Neighborhood operations generate each output pixel as a function

15   of several neighboring input pixels. Neighborhood size is often 3 x 3, 5 x 5, though it can be larger, smaller or shaped otherwise, in accord with requirements of a particular application.

Thresholding is an image enhancement technique for reducing the number of intensity, brightness or contrast levels in an image. It is typically used to convert a gray scale

20   image, with up to 256 gray levels, to a binary image, with just two levels (e.g., black and white). If a pixel intensity value exceeds a threshold (or is outside a threshold range), it is converted to a value that represents "white" (or potential defect); otherwise, it is converted to a value that represents "black" (or "background"). Threshold levels can be set at a fixed gray level for an image (level thresholding), or can be based on a variety of other measures, e.g.,

25   they can be set relative to an average gray scale level for a region (base line thresholding).

Thresholding is commonly used in machine vision systems to facilitate detection of defects. Prior art techniques, however, do not perform very well on non-woven materials, such as disposable diaper fabrics. When these materials are backlit and imaged at high

30   resolution, both embossing and normal variation in the material's "formation" (the structure of the material's fibers) can appear at the pixel level as small holes and/or thin spots amidst thicker, more solid regions. This makes it difficult to discern them from actual defects.

Traditionally, inspection system providers have solved the problem of inspecting such materials in one of two ways. They either image the materials at low camera resolutions, so minute variations in the materials are effectively lost, or they opto/mechanically defocus the camera lens, blurring the material variations to achieve somewhat the same effect. Both these

5   techniques result in poor image quality and, therefore, cannot be used in applications where high-resolution images must be displayed, e.g., for operator evaluation. Moreover, both result in a loss of valuable image data at the acquisition stage and, therefore, preclude further automated analysis.

10   In view of the foregoing, an object of the invention is to provide improved methods and apparatus for machine vision. A more particular object is to provide improved methods and apparatus for thresholding images.

A related aspect of the invention is to provide such methods and apparatus as facilitate

15   imaging and analysis of defects (or other features) in images.

A further aspect of the invention is to provide such methods and apparatus as facilitate the inspection of non-woven and other materials with high-frequency variations of intensity, brightness, color or contrast.

20

Yet another object of the invention is to provide such methods and apparatus as can be readily implemented at low cost with existing machine vision software and/or hardware.

Summary of the Invention

The foregoing are among the objects attained by the invention which provides, in one aspect, machine vision inspection methods that take an average (or other such statistical

5   measure) of pixel values within each of selected "neighborhoods" or groups of pixels within an image. The averages -- effectively, digitally "defocused" pixels – are thresholded and the results output for display or further analysis.

Such methods are advantageous because an originally acquired, high-resolution (non-

10   defocused) image can be preserved and processed in parallel with a neighborhood-based defocused and thresholded image. Systems employing these methods achieve the thresholding capability of traditional defocused systems, while providing clear, detailed, high-resolution images for display or other analysis. Such systems provide this dual capability using image data acquired from a single camera or camera array.

15

A method as described above can be used, for example, for inspection of webs of baby diaper "fabric." An arithmetic average is generated for each M x N unique (but overlapping) neighborhood of pixels in an image of the fabric. The averages, again, representing "defocused" pixels, are compared with a threshold (or threshold range) for the

20   fabric as a whole. Averages that exceed the thresholds are set to one value (e.g., "white," representing potential defect pixels); otherwise they are set to another value (e.g., "black," representing background). An image consisting of these thresholded, defocused values -- each positioned in accord with the location of its respective neighborhood in the original image -- can be displayed or used for inspection of the web.

25

Further aspects of the invention provide methods as described above in which an image consisting of the thresholded, defocused values is displayed superimposed on, or otherwise in conjunction with, the originally acquired image. Related aspects provide for display with image that is thresholded on a traditional per pixel basis.

30

Yet further aspects of the invention provide methods as described above in which the aforementioned threshold (or threshold range) is not fixed for the image as a whole but, rather, varies along the image. For example, the defocused pixel value for each neighborhood can be compared to a threshold based on an average pixel intensity for a larger region in

3

which that neighborhood resides. In the diaper fabric example above, such a method can be used to compensate for slowly changing variations in web color or brightness over an entire roll of web.

5   Still further aspects of the invention provide machine vision apparatus that operate in accord with the methods described above. An apparatus according to one such aspect includes a filter that generates an average pixel value of an M-column by N-row neighborhood. The filter includes a down delay memory that holds each pixel entering the filter ("new pixel value") for a specified period. A down accumulator having J storage

10   elements, where J is a number of columns in the acquired image, maintains a sum of N rows of pixel values for each of J corresponding columns of the image.

The filter further includes down accumulator logic that updates the down accumulators with each new pixel value received by the filter: adding the new pixel value to

15   a sum maintained by the down accumulator for the column with which the new pixel value is associated, subtracting therefrom a pixel value output by the down delay memory for that same column, and storing a result ("new down-sum") back into that down accumulator.

Moreover, the filter includes an M-element cross delay memory that holds each newly

20   calculated, per column down-sum for a specified period before outputting it to cross accumulator logic. That logic adds the newly calculated down-sum to a sum maintained in a cross accumulator, subtracts therefrom the "old" down-sum output by the cross delay memory, and stores the result ("new cross-sum") back to the cross accumulator. Newly calculated cross-sum values represent a sum of pixels values for the current M-column by N-

25   row rectangular neighborhood. Upon division by the product M * N, these are the defocused, neighborhood based pixel values discussed above.

These and other aspects of the invention are evident in the attached drawings and in the description and claims that follow.

30

Methods and apparatus according to the invention have utility in machine vision apparatus used in industry, research and other areas of pursuit. Such methods and apparatus facilitate the rapid and accurate inspection of non-woven and other materials or scenes with a high degree of color, contrast, intensity or brightness variation at the pixel level. In addition,

4

they facilitate low-cost simultaneous generation and/or use of corresponding unfiltered or per-pixel thresholded images.

Brief Description of the Drawings

A more complete understanding of the invention may be attained by reference to the drawings, in which

5

Fig. 1 is a schematic of machine vision system of the type used to practice the invention;

Fig. 2(a) shows a raw image of the type processed by the system of Fig. 1;

10

Fig. 2(b) shows the effect of conventional per-pixel thresholding on an image of the type shown in Fig. 2(a);

Fig. 2(c) shows the effect of digital defocusing and thresholding on a neighborhood

15 basis on an image of the type shown in Fig. 2(a);

Fig. 3 is a flow chart depicting operation of the system 10 of Fig. 1; and

Fig. 4 depicts operation of a preferred filter used in practice of the invention.

20

Detailed Description of the Illustrated Embodiment

Fig. 1 depicts a machine vision system 10 of the type with which the invention is practiced. The system includes an image acquisition device 12 that generates an image of an

5 object 14 under inspection. The illustrated device 12 is a linear camera of the type conventionally used in machine vision for the inspection of a web or other moving object, though, the invention is suitable for practice with any manner or type of image acquisition devices. Illustrated object 14 is shown as a web, e.g., a sheet diaper fabric, under manufacture and moving in front of backlight 18, all in the conventional manner. Though

10 particularly suited for inspection of non-woven or other materials or scenes with a high degree of intensity, color, contrast or brightness variation at the pixel (or other small-scale) level, it can be used for display and/or analysis of any acquired image.

Digital image data (or pixels) generated by the acquisition device 12 represent, in the

15 conventional manner, the image intensity (e.g., color, contrast or brightness) of each point in the field of view of the device 12. That digital image data is transmitted from capturing device 12 to image analysis system 20. This can be a conventional digital data processor, or a vision processing system of the type commercially available from the assignee hereof, Cognex Corporation, as outfitted and/or programmed in accord with the teachings hereof.

20 The image analysis system 20 may have one or more central processing units (CPU), memory units (Mem), input-output sections (I/O), and storage devices (Disk), all of the conventional type. Those skilled in the art will appreciate that, in addition to implementation on a programmable digital data processor, the methods and apparatus taught herein can be implemented in special purpose hardware.

25

Illustrated image analysis system 20 is coupled to a display device 26 of the conventional type and in the conventional manner. In the drawing, this is employed to illustrate one practice of the invention. An image acquired by device 12 is output to and displayed on device 26, e.g., with image magnification and/or color and contrast

30 enhancement, all in the conventional manner. Superimposed thereon is a form of the image processed in the manner described below to digitally defocus and threshold the pixels on a neighborhood basis.

In a preferred practice of the invention, the image that has been processed to digitally defocus and threshold the pixels on a neighborhood basis is routed for further automated image analysis. That image alone or, preferably, in combination with the originally acquired, high resolution, unfiltered image, can be used in such automated analysis for highly accurate,

5   automated detection of defects in a manner that avoids the misidentification of normal regions of the image as defective.

A further understanding of the illustrated system may be attained by reference to Figs. 2(a) – 2(c). Fig. 2(a) shows an image acquired from device 12 of web 14, with only

10   magnification and/or color and contrast enhancement. Region 28 identifies a clump or other defect in the web. Area 30 indicates a region of normal acceptable material formation.

Fig. 2(b) shows the effect of conventional per-pixel thresholding on the image of Fig. 2(a). As shown by the agglomeration of dark dots, this thresholding technique highlights the

15   defect in region 28. However, normal intensity variations in the region 30 of the web result in improper thresholding and highlighting of pixels there, as well as at other locations dispersed about the image. Though additional image processing (e.g., erosion) can be performed on threshold image of Fig. 2(b) in order to eliminate some of this false highlighting, that would also tend to de-emphasize the otherwise desirable highlighting in

20   region 28.

Figure 2(c) is an expanded view of the image on display device 26 and shows the effect of defocusing and thresholding on a neighborhood basis in accord with the teachings hereof. Again, an agglomeration of dark in region 28 reveals the defect there, yet, with more

25   emphasis than shown above. Moreover, unlike the conventional per-pixel technique, there is little or no highlighting in the region 30 or in the other non-defective regions of the web.

Comparing Figs. 2(a) – 2(c), those skilled in the art will appreciate that the dark pixels of Figs. 2(b) – 2(c) are superimposed over the original image (of Fig. 2(a)) to identify pixels

30   that have exceeded a threshold. Normally, such superposition would be displayed in color, e.g., red.

Those skilled in the art will appreciate that Figures 2(a) - 2(c) show just one application of the invention, and that in other applications defocusing and thresholding on a

neighborhood basis can result in more or less highlighting than conventional per-pixel techniques, e.g., depending on the intended application and how thresholds are set and used.

Fig. 3 is a diagram depicting operation of the system 10 of Fig. 1. In step 40, an
5   image captured by device 12 is passed to image analysis system 20 for processing. This step is performed in the conventional manner.

In step 42, the stream of pixels from the acquired image is passed through a filter that digitally defocuses them on a neighborhood basis. In the illustrated embodiment, the filter is
10   an arithmetic mean filter operating with a neighborhood (or window) size of 2x2, 2x3, 3x2, 3x3, 3x4, 4x3, 4x4, or any other size or shape (e.g., approximated circle) suitable to the inspection task at hand.

Preferably, the neighborhoods are rectangular and contain either a total pixel count
15   that is a factor of two or, alternately, have a width and height that are each a factor of two. This facilitates performing divisions required for averaging. The invention is not limited to arithmetic mean filters and may, instead, use filters that provide any statistical measure or low-pass filtering of pixel contrast, color, brightness, or intensity with the respective neighborhoods.
20

The result of the filtering step 42 is a stream of "defocused " pixels, each of which represents an average (or other statistical measure) of the neighborhood of pixels surrounding (and including) each pixel of the acquired image. In essence, the filtering step 42 has the effect of defocusing pixels in the acquired image, blurring them to the extent that small
25   variations are not noticeable. This does not deteriorate the sharpness of the image features larger than the filter window, yet it significantly attenuates noise, insignificant variations and features smaller than the neighborhood or window size. Details of a preferred filtering operation are discussed below and shown in Fig. 4.

30   In steps 44 and 46, the defocused pixels are thresholded. Thresholding step 44 involves comparison of each of the defocused pixels with a level threshold or range, set empirically or otherwise, for the entire acquired image. Defocused pixel values that exceed the thresholds are set to one value (e.g., "white," representing potential defect pixels); otherwise they are set to another value (e.g., "black," representing background). These

9

resultant values can, themselves, can be regarded as pixels having a binary value (e.g., black/white, 0/1, defect/background, etc.) .

Thresholding step 46 operates similarly, although it uses a base line threshold that
5   varies along the image. Determination of the threshold is accomplished in the conventional manner of the art. For example, in the illustrated embodiment, a running average of pixel intensities over a region – e.g., 128 x 128 pixels -- is maintained, e.g., utilizing accumulators of the type described below, or in software executing in the CPU of device 20. Those running averages are used in the thresholding step 46, as indicated by the arrows.
10

As noted, the illustrated embodiment utilizes two thresholding steps: level (step 44) and base line (step 46). Those skilled in the art will appreciate that one or both of these steps can be performed, depending upon application at hand. Moreover, it will appreciated that other thresholding steps can be performed instead or in addition. Further, as discussed below,
15   in embodiments where filtering step 42 does not include a division operation (e.g., pixel value sums for each neighborhood are not divided by the number of pixels in the neighborhood), the thresholding steps 44, 46 can employ threshold values that are scaled commensurately larger (e.g., the threshold values are multiplied by the number of pixels in the neighborhoods). Regardless of how performed, the result of the thresholding steps 44, 46
20   is a stream of binary pixel values that represent thresholding of the defocused pixels in each neighborhood of the acquired image.

In step 48, the binary pixels from the thresholding steps are output for display, e.g., in the manner shown in display 26 or Figure 1, and/or for further processing. In the either
25   event, those binary pixels can be combined with one another (e.g., via an OR operation), as well as with binary pixels from the conventional per-pixel thresholding operations 50, 52. Moreover, the results can be used to isolate and identify defects in the acquired image.

Concurrent with generation of the binary pixel values based on the defocused,
30   neighborhood-based pixel values, the method utilizes conventional techniques to threshold the stream of pixels from the acquired image on a per-pixel basis. See, steps 50, 52. These steps operate in the conventional manner known in the art, using thresholds set in the manner described in steps 44, 46, above (though these thresholds need not have the same values as

those used in the prior steps). Thresholded pixel values resulting from steps 50, 52 can be used as described above.

In a preferred embodiment, the stream of pixels from the acquired image are also
5   processed using "streak processing" thresholding techniques of the type known in the art. These are intended to discern streak-like defects occurring along the direction of motion of web 12 that might otherwise escape emphasis via the thresholding techniques performed in steps 42 - 52. Those skilled in the art will appreciate still other thresholding techniques may be utilized, and their results displayed and/or combined with the other threshold values as
10   above.

Hardware and/or software-based delay mechanisms of the type known in the art can be employed to align the acquired image and the aforementioned binary threshold images on a per pixel basis.
15

Fig. 4 depicts a preferred implementation of the filter whose operation is discussed above. In the drawing, the acquired image comprises rows (lines) of pixel data, each row containing J columns of pixels. Each pixel is an 8-bit value representing a gray level, where 0 is black and 255 is white. The image streams in continuously (i.e., it is not limited in the
20   number of rows it can contain), but for the sake of discussion, we can say it consists of K rows of data. The image is processed as a continuous stream of pixel data, with the pixels in each row following the pixels of the previous row.

An arithmetic mean filter calculates the average pixel value of every unique M-
25   column by N-row rectangular sub-image (neighborhood) within a given image (including overlapping sub-images). In effect, the pixel values in each M by N sub-image are added together, then that resultant sum is divided by the number of pixels in the sub-image (the product (M * N)). The result is a value that is the average of the pixel values in the sub-image.
30

If these average values are arranged in columns and rows corresponding to the relative position of their respective sub-images in the original image, the average values themselves make up an image, smaller than the original image ((J-M+1) columns, (K-N+1) rows).

In a preferred embodiment, running averages are used so that each pixel in the original image is referenced the fewest times possible.

A pixel stream enters the arithmetic mean filter. A down delay is a memory (e.g.,
5 FIFO or RAM) that holds each pixel value entering it (or stored to it) for N rows, before outputting (or accessing) the same pixel value. In other words, as each pixel is stored to the down delay, it is held, then referenced (J * N) pixels later. We can call each pixel value entering the down delay a "new" pixel value, and each pixel exiting the down delay an "old" pixel value.
10

A down accumulator is a memory containing one storage element for each of the J columns in the image. Each memory element has sufficient data bits to hold the maximum sum of N image pixel values (the maximum sum of an N-row high column of pixel values in the sub-image). The value stored for each column is called a down-sum, and represents the
15 sum of N rows of image pixel values in that column.

During operation, for each column, the down accumulator logic takes an existing down-sum from the down accumulator, subtracts off the old (delayed) image pixel value for the column, adds in the new image pixel value for the column, and stores the new down-sum
20 back in the down accumulator. For each new incoming row in the image, new per-column down-sums are calculated, using this running average method.

To initialize the down accumulator upon startup, the process takes place as describe above, with the accumulator logic forcing the existing down-sum to zero for every column in
25 the first row, and forcing the old image pixel value to zero for every column in the first N rows. This allows the down accumulator to initialize by summing together the first N rows of image pixels in each column before performing a running average. Note that the down-sums in the down accumulator are not valid for the first (N-1) rows.

30   As each newly updated down-sum is being written back to the down accumulator, that down-sum is simultaneously passed to the cross delay. A cross delay is a memory (e.g., FIFO or RAM) that holds each down-sum entering it (stored to it) for M columns, before outputting (or accessing) the same down-sum. In other words, as each down-sum is stored to the cross delay, it is held, then referenced M columns later. We can call each down-sum

entering the cross delay a "new" down-sum, and each pixel exiting the cross delay an "old" down-sum.

5   A cross accumulator is a single memory element with sufficient data bits to hold the maximum sum of (M * N) image pixel values (the maximum sum of the pixel values in the sub-image). The value in the cross accumulator is called a cross-sum, and represents the sum of pixel values in an M column by N row sub-image.

10   During operation, the cross accumulator logic takes the existing cross-sum from the cross accumulator, subtracts off the old (delayed) down-sum, adds in the new incoming down-sum, and stores the new cross-sum back in the cross accumulator. For each new incoming per-column down-sum in the image, a new cross-sum is calculated, using this running average method.

15   To initialize the cross accumulator upon startup, the process takes place as describe above, with the accumulator logic forcing the existing cross-sum to zero for the first new down-sum in each row, and forcing the old down-sum to zero for the first M new down-sums in each row. This allows the cross accumulator to initialize by summing together the first M columns of down-sums in each row before performing a running average. Note that the 20   cross-sum in the cross accumulator is not valid for the first (M-1) new down-sums in a row.

      As each newly updated cross-sum is written back to the cross accumulator, it is simultaneously passed to a divider. The cross-sum is the sum of all pixels in an M column by N row sub-image. The divider divides the cross-sum value by the product (M * N). The 25   result is the average pixel value in each M x N sub-image. This average can then be routed for thresholding or additional processing as described above.

      There are variants that minimize and simplify this implementation. Arbitrary division can results in complex implementation or slow performance . If the total number of pixels in 30   the averaging sub-image is restricted to being a power of two, a simple high-order bit sub-selection (shifting) can be used to effect the division, resulting in a slight loss of accuracy by truncation. Another alternative is to limit each dimension (rows and columns) of the sub-image to a power of two. In this case, the down-sums can be shifted (divided by the sub-image height N) prior to being passed to the cross delay and cross accumulator logic. Each

cross-sum can be shifted (divided by the sub-image width M) prior to being passed on for further processing. Additional inaccuracy is introduced by taking an "average of averages", but in some applications this may be acceptable. Lastly, in some cases, it may not be necessary to divide the cross-sums at all. Some thresholding and other processing can work

5   on an accumulated sum rather than an average. This can often be accomplished by scaling threshold to the magnitude of the sub-image size. For example, threshold values operating on 7 x 5 sub-images may be scaled up by a factor of 35 to operate on the sub-image sum rather than the sub-image average.

10   Discussed above are systems and methods meeting the desired objects. It will be appreciated that the illustrated embodiments are merely examples of the inventions, and that other embodiments incorporating changes therein may fall within the scope of the invention, of which we claim: